The Effectiveness of Marketing Strategies in Social Media: Evidence from Promotional Events

Panagiotis Adamopoulos
padamopo@stern.nyu.edu

Vilma Todri
vtodri@stern.nyu.edu

Department of Information, Operations, and Management Sciences
Leonard N. Stern School of Business, New York University

ABSTRACT

This paper studies a novel social media venture and seeks to understand the effectiveness of marketing strategies in social media platforms by evaluating their impact on participating brands and organizations. We use a real-world data set and employ a promising research approach combining econometric with predictive modeling techniques in a causal estimation framework that allows for more accurate counterfactuals. Based on the results of the presented analysis and focusing on the long-term business value of marketing strategies in social media, we find that promotional events leveraging implicit or explicit advocacy in social media platforms result in significant abnormal returns for the participating brand, in terms of expanding the social media fan base of the firm. The effect is also economically significant as it corresponds to an increase of several thousand additional new followers per day for an average size brand. We also precisely quantify the impact of various promotion characteristics and demonstrate what types of promotions are more effective and for which brands, while suggesting specific tactical strategies. For instance, despite the competition for consumers' attention, brands and marketers should broadcast marketing messages on social networks during the time of peak usage in order to maximize their returns. Overall, we provide actionable insights with major implications for firms and social media platforms and contribute to the related literature as we discover new rich findings enabled by the employed causal estimation framework.

Categories and Subject Descriptors

K.4.4 [**Electronic Commerce**]: Online Shopping; J.4 [**Social and Behavioral Sciences**]: Economics

General Terms

Economics; Experimentation; Measurement

Keywords

Business Analytics; Causal Inference; Event Study; Counterfactual; Social Media; Business Value; Social Commerce

KDD'15, August 10-13, 2015, Sydney, NSW, Australia.
 ISBN 978-1-4503-3664-2/15/08 ...$15.00.
DOI: http://dx.doi.org/10.1145/2783258.2788597.

1. INTRODUCTION

Social media are fundamentally changing the way we communicate, collaborate, and consume. They represent one of the most transformative impacts of information technology on business [3] as they drastically change how consumers and firms interact. Harnessing the opportunities that social media generate, marketers nowadays are seeking novel ways to create or further increase the consumers' connection and engagement with their brand. As consumers spend an increasing amount of their time online and the percentage of adults using social media significantly rises, companies invest a growing amount of their marketing budget towards online and social media advertising, invent new ways to establish strong connections with their customers into the online world, and leverage their social connections. Therefore, companies nowadays increasingly compete for consumers' attention and engagement with their brand in the social media space. Content generation, nurture of positive online word-of-mouth (WOM), and utilization of social links among customers are some quintessential effective means of non-paid advertising for companies to spread their message in real time and generate leads associated with key marketing objectives.

Social commerce is a representative example of how e-commerce can leverage the social connections between users to generate leads for a business [31]. The term "social commerce" was first introduced in 2005 by Yahoo! [27] and initially referred to a feature that allowed users to review products on the Yahoo! shopping platform. During the next years, the scope of the term "social commerce" expanded in various directions incorporating peer recommendations, shared shopping lists, product referrals, coupon sharing, team-buying (i.e., a group of customers gathering together in order to bargain with merchants), "deal-of-the-day" websites (i.e., e-commerce websites where a minimum number of purchases should be reached in order an offer to be activated), network-based marketing and integration of online merchandisers with social network platforms, and company-controlled online communities [32]. Towards this direction, American Express Company (Amex) recently pioneered a social commerce venture on the social media platform of Twitter, Inc. based on which a user can purchase a product or claim a promotional offer from a participating brand by sending a short text message, called "tweet", with a designated keyword (i.e., hashtag). The service has already been deployed in large scale offering to users credit of $15 million just in April 2014 [1] and $100 million in total

(Sept 2014).[1] The social commerce initiative of Amex on the social network of Twitter builds on concepts related to the previous social commerce ventures but also exhibits some unique features. For instance, in this setting all the discount claims and transactions are by default visible to the users of the platform, each purchase and offer takes place within the platform itself using only the core feature of the platform (i.e., designated hashtags), the promotions correspond to both online shops and brick-and-mortar businesses, both physical and electronic products are available, and the products and offers are delivered to the user that completed the transaction instead of other users (e.g., Facebook gifts).

The aforementioned unique features of this new venture and the simplicity in completing purchases and claiming offers generate tremendous potentials as they can fundamentally change the way consumers shop online and interact with brands [31]. Wrapping the entire purchase (and discount claim) process into a hashtag, the service makes the transactions fast and easy for the consumers. At the same time, the proposed service capitalizes the unique opportunity for social media to generate direct return on investment. Thus, the *short-term effects* of such an offering seem straightforward for both the consumers, who save money on their purchases, and the marketers, who generate buzz [13] and sales [31] around their brand.

However, this paper assumes the point of view of social brands and organizations and attempts to focus on the *long-term effects* of such social media marketing strategies enabled by the unique salient features of the social medium. Amex essentially turns every single purchase or offer claim into an advertisement as the social neighbors of the customer get exposed to the public advocacy and endorsement of the brand or the product from a fellow friend. The service is an exemplary model of how a highly engaging social media campaign could generate commitment from the consumers and reinforce their loyalty to the brand. Therefore, we attempt to elucidate the role of marketing strategies and promotional events in increasing the user fan base for the participating brand or organization. Any change in the size of the social media fan base of a company is a visible important indicator of firm performance, since a user following a brand on Twitter can be considered as implicit word-of-mouth and recommendation. Besides, companies and organizations are particularly interested into attracting more followers for their social media profiles and capitalize on every opportunity to do so as in the long-run they can more easily generate brand awareness, create engagement, and spur online WOM. Furthermore, a larger fan base might further enhance the company's ability to harness the power and effectiveness of social media in the future as well [31]. According to the hierarchy of effects in marketing, such key marketing objectives are directly linked to increased sales [30] and it has been shown that the corresponding social media-based metrics are significant leading indicators of firm equity value [21].

In this paper, we seek to understand the long-term effectiveness of social media marketing strategies that integrate public advocacy and endorsement features by examining their impact on the image of the participating brands and organizations, specifically their fan base. Focusing on the long-term business value, we employ an event study methodology combining econometric with predictive modeling techniques in a causal estimation framework that allows for more accurate counterfactuals and use a real-world data set consisting of *all* the promotional events by Amex and other participating brands on Twitter, more than 70 million brand-related tweets, and detailed information for more than 3 million users. We find that marketing strategies incorporating implicit or explicit advocacy on social media platforms result in significant abnormal returns for the participating brand, in terms of additional new followers. We also precisely quantify the impact of the various promotion characteristics (e.g., discount, price, flexibility, expected benefit) and demonstrate what types of promotions are more effective towards expanding a brand's social media fan base. Besides, we illustrate for which brands such promotional events are more effective and suggest specific tactical strategies. Overall, we contribute to the related literature discovering new rich findings and provide actionable insights with demonstrable implications for both marketers and social media platforms.

In Section 2 of this paper, we develop the hypotheses pertaining to our research questions and discuss the related work. Then, in Section 3, we describe the particular e-commerce business model as well as the data generating process and, in Section 4, we describe in detail the employed methodology for causal estimation in combination with the deployed data mining techniques. Section 5 presents our empirical results and the real-world value of our discoveries, Section 6 discusses the additional robustness tests we conducted, and Section 7 describes the falsification checks ("placebo" studies) we deployed in order to verify our causal estimates. Finally, Section 8 concludes this paper summarizing our findings as well as the limitations of this work.

2. RELATED WORK AND HYPOTHESES DEVELOPMENT

Employing the hashtag feature of Twitter, American Express effectively transforms the specific social media platform into a social commerce platform, where consumers can directly buy products or claim promotions with the easiness of a tweet and automatically spread the word about the offers. This novel emergent type of e-commerce and user engagement is related to several streams of research from different fields, including Computer Science, Information Systems, Marketing, and Economics. These streams of research include network-based marketing, mobile coupons, real-time marketing, and word-of-mouth (WOM). We review the extant literature to summarize previous work and develop our research hypotheses. Additionally, we discuss the related work regarding the employed causal estimation framework.

Network-based marketing refers to a collection of marketing techniques that take advantage of links between consumers to increase sales and create buzz. Different modes of network-based marketing have been identified in prior literature. For instance, "network targeting" refers to a firm marketing to prior purchasers' social-network neighbors, possibly without any action by the customers, whereas "advocacy" involves particular individuals becoming either vocal advocates (i.e., explicit advocacy) for a product or service recommending it to their friends and acquaintances or implicit advocates through their own actions (e.g., adoption of a specific service or product) [14]. In the particular setting of Twitter and Amex, consumers automatically reveal their

[1] `https://twitter.com/AmericanExpress/status/511591536082296832`

purchases and, thus, contribute to the implicit advocacy. Consumers may additionally choose to personalize the tweet messages they share with their social neighbors and, thus, become explicit advocates. Hence, Amex and its partners using complimentary modes of network-based marketing as an integral part of the purchase process aim at taking full advantage of the social links between customers in order to drive awareness, sales, and engagement. Thus, as the customers spread the word about an offering and advocate the corresponding brand through their product endorsements, more social media users get exposed to the promoting brand and its activities and through a variety of influence processes, including those that raise awareness as well as those that persuade individuals to change their expectations [2], might decide to become social media followers of the particular brand in order to keep up to date with the company's products, services and offerings. On the other hand, promotional events might have a negative impact since such promotions may lead to inferences about the value and quality of the brand and its products. In particular, consumers, especially those that are not familiar with the brand, might interpret these promotions as a signal that the specific product is of low quality or that the firm is in trouble [10, 29]. However, in the online world, thanks to the advancements of modern information systems as well as the valence and volume of online reviews, consumers have easier access to higher quality information, which in combination with the lower search costs improves their ability to compare different brands and products [4], and thus the potential influence of such negative signals is reduced. Besides, because of the increased price competition and the more frequent price adjustments online [6] as well as the increased price sensitivity of consumers [7], such promotions are also less likely to be perceived as negative signals. Hence, promotional events on social media are less likely to have a negative impact. Therefore, we propose that:

$H1$: *A social media marketing strategy incorporating features of implicit and/or explicit advocacy is likely to attract new followers for the participating brand or organization.*

Besides, the current fan base of a brand might affect the potential of social media marketing strategies to attract new followers through online WOM. Research suggests that both volume and valence of WOM have a significant impact [8] while some prior empirical studies indicate that the volume of WOM is more effective than the valence through the awareness effect [20]. Especially in online social media and networks, the large number of users offers the potential for online WOM at much larger scale, causing online social networks to render online WOM more convenient than traditional WOM [9, 35]. Since new customer acquisition is crucial for businesses and given the aforementioned importance of social networks, social media platforms have attracted the attention of many companies who have a perennial interest in leveraging social relationships to extend their customer base [28]. Nonetheless, the current network fan base could either amplify or attenuate the effectiveness of marketing strategies on attracting new followers. For instance, despite the awareness effect of such social media marketing strategies, it might be more difficult for a company with a larger existing fan base to attract new fans due to saturation effects since it is more likely for a user to have already made a conscious decision in the past whether to follow the particular brand or not. On the other hand, a large fan base might positively affect customers' decision to follow a brand through a social influence bias [24] contributing to a "rich-get-richer" effect. Therefore, we hypothesize that:

$H2a$: *A social media marketing strategy from a brand with a smaller existing fan base is less likely to attract more followers for the participating brand or organization.*

$H2b$: *A social media marketing strategy from a brand with a smaller existing fan base is more likely to attract more followers for the participating brand or organization.*

Moreover, Twitter as a social medium is closely aligned with the trends towards mobile [5] and real-time marketing. Twitter is currently one of the few social platforms that has already captured the "in-the-moment" context and this allows marketers not only to propagate their messages in a timely manner but also to capitalize the context and the mindset in which the consumer is more likely to be receptive and responsive to an advertiser's message [33]. Hence, advertisers who would like to embrace the real-time opportunities of Twitter would be interested into choosing the correct time of promoting such events on the Twitter platform in order to maximize the return on engagement/investment. In particular, someone might argue that the time of peak usage would increase the chances of exposure to a greater audience and thus would attract more followers. However, it could be also the case that at the time of peak usage propagating a promotional message is very competitive and so less effective in attracting consumers' engagement and attention. Therefore, we hypothesize the following:

$H3a$: *A promotional message that is broadcasted at the time of peak usage of the social media platform is more likely to attract more followers for the participating brand or organization.*

$H3b$: *A promotional message that is broadcasted at the time of peak usage of the social media platform is less likely to attract more followers for the participating brand or organization.*

Furthermore, the ability of a marketer to attract new followers will depend on the perceived benefits of the promotional event as well as the expectations users form about future events. Dickinger and Kleijnen [11] investigate consumers' intentions to redeem mobile coupons since coupon redemption rates are important drivers of sale increases, profits, and market performance [18, 25]. They suggest that the economic benefit is a significant determinant of the coupon usefulness and attractiveness and, consequently, we theorize that, *in a social commerce setting, a promotional event with a larger percentage discount is more likely to attract more followers for the participating brand.* Besides, the ability to claim an offer multiple times allows for additional monetary savings and would be likely to generate more followers who benefit from accumulated savings. Therefore, *a promotional event with an offer that can be claimed multiple times by a single customer is more likely to attract more followers for the participating firm.* While higher savings would be likely to attract more followers, the minimum required amount a customer is required to spend also comes into play. Hence, *a promotional event requiring* *a larger amount (USD) of minimum spend is less likely to attract more followers for the participating brand.* Additionally, Molitor et al. [23] show that the closer the consumers are to the physical store offering the coupon, the more likely they are to download the mobile coupons. Thus, *in a social commerce setting, a promotional event that is valid only in brick-and-mortar stores and not online is hypothesized to be less likely to attract more followers for the participating brand.* Based on the above discussion, *a promotional event with higher expected benefits for consumers is more likely to attract more followers for the brand* and hence we measure the corresponding effects.

Moreover, our work is closely related to prior research in e-commerce. Studying the consumer behavior in this modern social commerce service, in [31] we elucidate the factors that drive and affect the consumers' decision to adopt this novel service and make a purchase that will be disclosed automatically to the social network. In particular, employing both econometric and predictive models, we study how the characteristics of the user and her social network affect the decision to make a purchase and engage into WOM at the same time, in an attempt to better understand how firms can conduct business through social commerce and spur online conversations. Based on the empirical results, we find that various user characteristics, the brand loyalty and trust of the users, and their familiarity with the platform have significant effects on the likelihood of adoption and social purchases. Additionally, we find that the economic behavior of a user's immediate social network as well as the personalization and the valence of recommendations from the social neighbors of the user have also significant impact on her/his decisions to make such purchases. Besides, observing the WOM episodes, the breadth of their dissemination, and the valence of the recommendations, we are able to study the distinguishing characteristics of the disseminators that are associated with successful post-purchases (i.e., after the transaction of the disseminator has been broadcasted to the network) of their neighbors either due to awareness or influence effects. The derived models provide specific guidelines for marketers to orchestrate WOM in social networks.

As far as the *methodological framework* employed in this paper is considered, event studies represent a relatively underexplored but promising research approach which looks at abnormal returns and reactions to specific events or happenings and measures directly the changes in the quantity of interest. Events studies have been mainly used in economic and finance research where they provide a powerful setting to examine the informativeness and impact of an event, such as investments and acquisitions, on the stock price of a firm. However, at their most general level, event studies do not necessarily include or require stock market information. Instead, there could be a relationship between an event and *any* dependent variable. Thus, the applicability of event studies is not limited to economic and financial research. Event studies have been used, for example, to identify specific factors that influence the outcomes of information technology (IT) investments by publicly traded firms. In recent years, the focus has been extended beyond IT investments to other issues, such as security incidents, IT outsourcing initiatives, e-commerce investments, and standardization projects [26]. Nonetheless, this paper is the first to combine data mining techniques with the econometric methodology of event studies in a causal framework in order to identify the true effect of social media marketing strategies on the user base of brands and organizations. Comparing event studies with other methodologies for estimating average treatment effects (ATE), such as the difference-in-differences (DiD) method or model free approaches, event studies are not limited to only two time periods (i.e., pre-intervention and post-intervention), they do not assume that the event (or treatment) is exactly the same for all treated members, that it has the same effect in every time period, nor that it is static and has no duration, but event studies can include flexible leads and lags. Besides, among other differences, event studies do not necessarily need control groups that consist of the same untreated individual members over time and hence do not suffer from a possible violation of the parallel trend assumption, even though event studies are flexible enough and can include such control groups or "market" observations, if preferable. Similarly, event studies can include synthetic controls (regardless of the scale of the dependent variable across observations and the sign of the correlation between the treated and control units) but they do not need to match individual observations on observed characteristics while assuming that they are also similar on unobserved traits, as in propensity score matching (PSM) techniques. Hence, the proposed methodology can contribute to the formation of synthetic control groups that do not violate the established conditions for causal inference from observation data. In addition, tapping into the advances of predictive modeling and leveraging (big) data of high veracity and variety, the proposed methodology also allows for more accurate counterfactuals, apart from more appropriate control groups. Thus, event studies can be naturally combined with data mining and machine learning approaches for accurate causal inference. In Section 4, we describe in detail the proposed integration of event studies with data mining and machine learning techniques as well as the corresponding assumptions for causal inference.

3. THE BUSINESS MODEL AND DATA GENERATING PROCESS

The data generating process of the specific social commerce case we study differs significantly from traditional online commerce processes. In particular, American Express announces the list of participating brands and organizations and the corresponding promotional sales and offers with their respective terms. The featured promotional products and offers belong to a wide variety of categories, including retail, travel, entertainment, etc.

Figure 1: Amex announcing a promotion on Twitter.

A typical announcement for such a promotional offer is the following: *"Tweet #AmexWholeFoods, get $10 back 1x on $75+ purch at Whole Foods w/synced Amex Card! (RegLtd, Exp 6/30) Terms: http://amex.co/YQTj4Y"*, as shown in Figure 1. Respectively, a typical promotional sale announcement is the following: *"Get Xbox 360 4GB for $179.99+tax w/synced Amex Card. Tweet #BuyXbox360Bundle to start purchase! QtyLtd Exp 3/3 Terms http://amex.co/W4X7xa"*.

Then, consumers who are interested into making a purchase or claiming an offer must have a Twitter account and sync their Amex account with Twitter through an easy opt-in process. Once Amex announces the offers, users can purchase the products or claim the offers by posting a tweet (a message of maximum 140 characters) with the designated hashtag or by re-tweeting the specific announcement. Such a tweet is publicly posted on the Twitter profile of the user (private accounts were not eligible to participate in the program) and her/his social network friends will automatically receive the tweet on their own timelines. Amex tracks these tweets that use the designated hashtag in the social network and matches them to the desired product or offer. After the tweet is automatically processed, a reply is sent to the user verifying the terms of the offer or asking to confirm the purchase within fifteen minutes. In the case of a product purchase, once the transaction is confirmed by the user within the designated timeframe, Amex bills the customers or credits the promotional amount in a following eligible purchase, respectively. We should note that a user is not required to follow neither Amex nor the participating brand in order to purchase a product or claim an offer.

Our data set contains *all* the promotional events available by Amex and the participating brands on Twitter. For each promotional event, the available information in our database and the derived variables include, among others, the percentage discount of the promotional event, the minimum amount a user is required to spend in order to be eligible for the promotion, whether the promotion is available only for in-store purchases, whether the promotional event corresponds to a product purchase or a discount offer, whether the promotion can be claimed multiple times from a single user, the number of followers of the participating brand when the announcement of the promotional event was sent, and whether the offer was announced between 12:00pm and 4:00pm, which is the peak of user activity on Twitter platform [34]. The data set spans all the promotional offers and sales that took place since the program was launched, March 2012, until the end of January 2015. Table 1 summarizes the variables of interest and shows the corresponding descriptive statistics computed over all the observations in our data set.

Additionally, using the Twitter API we have access to the daily number of followers, friends, and statuses, since 2011 (or the date the corresponding account was created and verified by Twitter) and for each calendar day, for all the verified accounts on Twitter and all the official brand accounts. Finally, for each brand we have also collected the corresponding economic and market sectors in which it operates as well as their user social network. In summary, we have data on 288 promotional events from 189 brands as well as daily observations for more than 22,000 brands across various industries.

4. EVENT MODEL AND EMPIRICAL ANALYSIS OF MARKETING STRATEGIES

Typical event studies first require identifying the event of interest (e.g., announcement of a merger between two business entities). After the specific event is defined, the period of time over which the corresponding quantity of interest (e.g., stock price of the firm) is adjusted is determined. Then, the quantity of interest is observed and the changes in response to the event, beyond the "normal" or expected level in the absence of the event, are examined to determine the extent to which the event changes the quantity of interest [17]. Nevertheless, event studies are not limited to economic and finance research and do not necessarily include or require stock market information but there could be any relationship between an event and a dependent variable, without the need to presume market efficiency as in stock market studies; however, the motivation and the theories used to generate expectations are likely to differ across disciplines [17]. In such settings, event studies can be a superior framework to various alternative approaches in investigating the response to particular happenings.

In our setting, the main identifying assumption is the absence of confounding events during the estimation and event windows; we describe in detail in the following page how confounding events can be identified employing automated and / or manual processes. Besides, our models for generating expectations (or "normal" returns) are based on data mining techniques and methods following the corresponding machine learning and data mining principles to accurately predict the quantity of interest and avoid problems of over-fitting and data leakage. More formally, the conducted event study assumes that i) a set of features, which is predictive of the outcome variable in the pre-intervention period, would have predicted the outcome also in the absence of the event of interest, and that ii) this set of predictive features has not been affected by the intervention. Even though all causal inference methods require some key model assumptions, the aforementioned assumptions for the identification of the causal impact of the events of interest using the described approach with observational data are less strict and much more realistic compared to the assumptions of the alternatives approaches described in Section 2.

In the following paragraphs we describe in detail the deployed approach following the outline that MacKinlay described in the context of economics and finance [22]. In particular, an event study methodology involves the following steps: (i) identification of the *event of interest*; (ii) definition of the *event window*; (iii) selection of the *sample* set of firms to be included in the analysis; (iv) prediction of the *"normal" return* during the event window in the absence of the event; (v) estimation of the *"abnormal" return* within the event window, where the abnormal return is defined as the difference between the actual and predicted returns without the event occurring; and (vi) testing whether the abnormal return is statistically different from zero.

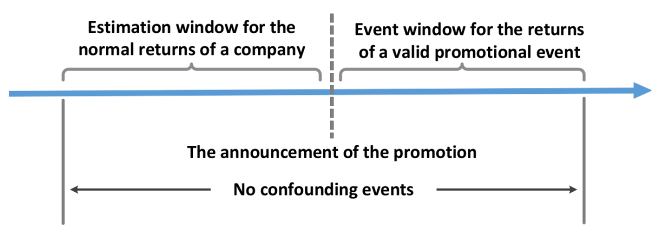

Figure 2: Graphical representation of the estimation window, event, and event window of a promotional event.

In the particular setting of social commerce, participating firms and American Express make announcements on Twitter about specific marketing promotions. Such announcements serve to inform the users about the corresponding terms of each promotion and, for this study, they constitute

Table 1. Variables and Descriptive Statistics

Variable	Description	Mean	Std. Dev	Min	Max
# Followers	The number of followers of the participating brand on Twitter	125,658.60	389,098.70	952	6,864,324
% Discount	The promotional discount	0.2487	0.1700	0.1	1
$ Minimum Price	The minimum amount (USD) a user is required to spend to be eligible for the promotion	126.61	213.10	0	2,500
In-Store Purchase	Whether the promotion is available only for in-store purchases	0.6684	0.4708	0	1
"Pay-By-Tweet"	Whether a user can purchase the promotional product using the 'Pay-By-Tweet' service	0.0320	0.1664	0	1
Unlimited Use	Whether the promotion can be claimed multiple times by a user	0.0285	0.1664	0	1
Social Network Peak Time	Whether the promotion was initially announced between 12:00pm and 4:00pm	0.3656	0.4816	0	1

the *events of interest*. Figure 2 shows the timing of an event relative to the estimation and event windows. If the promotion (i.e., event) is attractive to social media users, a positive response to this promotion is expected from their side. In the short-run, the company benefits from the generation of sale leads and the buzz created around its brand [31]. A central tenet of our approach though is that in the long-run this potential positive response can be translated into an abnormal increase of the user base of the participating brand. On the other hand, if the promotion is unattractive to the users, a negative impact on the user base can occur.

The *"event window"* indicates the number of days before and after the announcement date over which the abnormal returns are accumulated. An event window is typically denoted $(-x, +y)$, where x is the number of days before the announcement day and y is the number of days afterwards, while the announcement day is typically denoted as "day 0". Including days before the announcement captures information leaks, either from the press or internal users [17]. In order to capture such information leaks in this setting, we consider as the starting point of the event window for each promotional event the earliest time of the actual announcement of the promotion or the first message claiming the specific offer, if information leakage occurred; in some cases the first "response" to a promotion is observed a few hours before the actual announcement of the promotion because of information leakage via the employees of Amex or the participating companies. Besides, we narrow the width of the event window to five days refining the effect as precisely as possible and accurately estimating -a lower bound of- the causal impact of the event. This design choice is common in various settings and is motivated by the potential presence of confounding effects over a wide event window. The use of shorter windows reduces the potential for a confounding event to interfere as well as limits the impact of other events on the event of concern. If the event under study affects the dependent variable for a longer period than the selected event window and in the absence of confounding effects after the selected event window and during the influence of the particular event, we underestimate the total causal impact of the event but the results can still be interpreted as the causal impact of the event during the $x + y$ first days or, more generally, as a lower bound of the causal effect. These design choices are also consistent with the suggestions of Konchitchki and O'Leary [17] for effectively designing event studies.

Then, the *sample of firms* is chosen based on the particular event of interest and a combination of both automated and manual techniques. The set of candidate firms consists of all the participants that implemented a social media marketing strategy and had a promotional event on Twitter. The final sample though includes only the firms for which a confounding effect was not observed during the event window and the estimation window for the observations used in estimating normal returns. For instance, one of the candidate firms announced a social media marketing promotion during the time period when it sponsored a major festival and thus we excluded the firm from the final sample, since the change in the dependent variable could not be untangled from the confounding event and attributed to the particular social media marketing promotion. Eliminating such social media marketing promotions that may be tainted by another event or a set of events is an important step towards precisely estimating the causal effect of such promotions. The confounding events during both the event window and the estimation window were independently identified by two appointed graduate students and correspond to a small number of observations. Additionally, the LexisNexis database [19] was also employed in order to identify confounding effects based on full-text articles in newspapers, magazines, and journals from the U.S. and around the world, as in [15].

Besides, the significance of the impact of the event during the event window is usually assessed relative to what is referred to as the *estimation window* (or normal return period) over which a model of the normal returns as well as the variance of abnormal returns are estimated. The observations used to learn the model of normal returns for the counterfactual market response correspond to the specific brand and/or a set of similar companies, the control group, and other predictor variables; the specific set of control time series is described in detail in the following paragraphs. The data mining model learned during the estimation window is used to predict normal returns over the event window (i.e., expected returns at the absence of the event), which are then compared to the abnormal returns. The length of the estimation window is selected to be 120 days, ending immediately prior to the event window.

After estimating both the normal and abnormal (i.e., actual) returns over the event window, we focus on examining

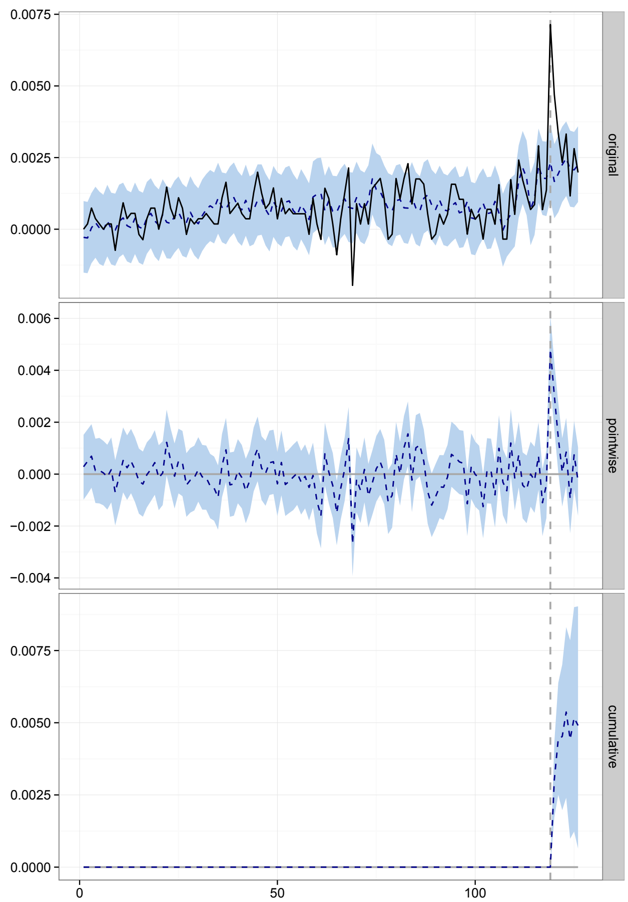

Figure 3: Effect of a promotional event on the number of followers of a participating brand.

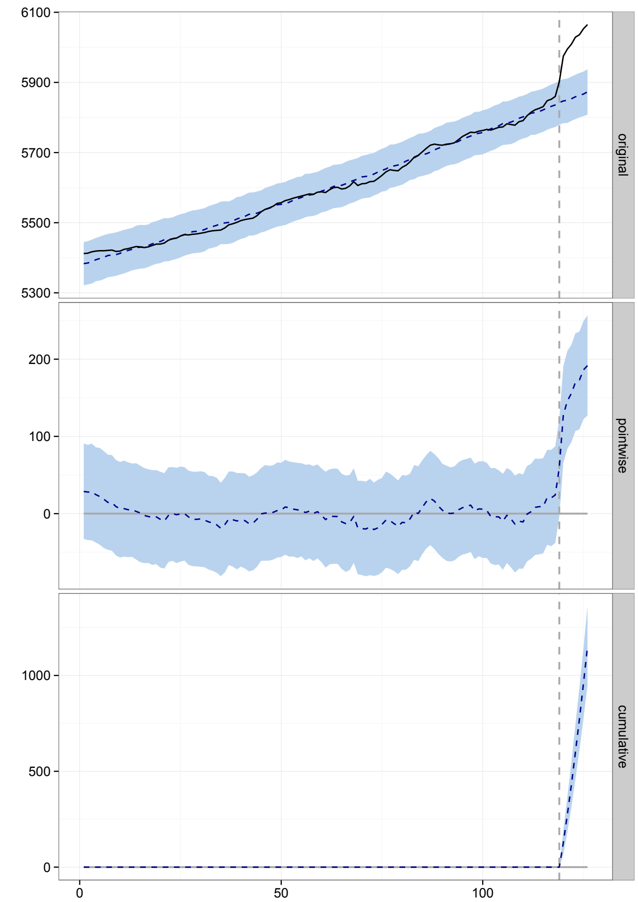

Figure 4: Effect of a promotional event on the number of followers of a participating brand.

the cumulative abnormal return as follows:

$$CAR_i = \sum_{t=-x}^{+y} AR_{it}$$

where $AR_{it} = R_{it} - \mathrm{E}(R_{it})$ is the point-wise difference between the actual observed outcome and the predicted counterfactual at the absence of an intervention, R_{it} the actual return of firm i on day t, $\mathrm{E}(R_{it})$ the corresponding expected return for the same firm and day at the absence of the event (i.e., based on the predictive model learned during the estimation window), and $t \in (-x, y)$. In particular, the predictive model of normal (expected) returns for each firm is

$$R_{it} = y_{it} = \alpha_i + \beta_i \mathbf{x}_{it} + \gamma_i^b \sum_{j \in B} w_{jt}^b + \gamma_i^f \sum_{j \in F} w_{jt}^f + \gamma_i^s \sum_{j \in S} w_{jt}^s + \delta_i s_{it} + \zeta_i z_t + \epsilon_{it}$$

where y_{it} is the natural logarithm of the ratio of number of followers in day t over the number of followers the previous day, $\mathbf{x}_{it}$ is a vector that contains dummy variables about the day of the week and whether there was a bank holiday, w_{jt}^b is the natural logarithm of the ratio of number of followers of brand j in day t over the number of followers at $t-1$ for B firms and organizations that have verified accounts on the social network, w_{jt}^f is correspondingly the logarithm of the ratio of followers for F similar firms (based on common followers), w_{jt}^f is the same metric for S similar firms (based on whether they belong to the same industry and economic sector), s_{it} the number of statuses (i.e., messages) posted on the social network by brand i at time period t, and z_t captures any time trend of the change in performance. These predictive models are then used to predict normal performance during the event window. The aforementioned predictor timeseries and the corresponding control firms of each brand and organization included in the above data mining model were selected based on i) all the officially verified accounts on the social network that correspond to a brand, ii) the relative number of common Twitter followers using the Jaccard coefficient and setting $F = 10$, and iii) all the verified firm-related accounts which belong to the same industry and economic sector.

The use of the specific formulation of the dependent variable allows the comparison of results across brands and better fits the large range of the quantity of interest; alternative formulations are discussed as robustness checks. Figure 3 shows for a specific firm the observed time series and the predicted counterfactual, the abnormal returns per time period, and the cumulative abnormal returns. For illustrative purposes, in Figure 4 we use as dependent variable the num-

Table 2. Cumulative Abnormal Returns Across All Events

Variable	Coefficient		Robust Std. Err.	t-statistic
Cumulative Abnormal Returns	0.0019596	***	0.00058	3.37

Robust standard error are reported
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3. Model of Abnormal Returns

Variable	Coefficient		Robust Std. Err.	t-statistic
Constant	0.0059222	***	0.0014266	4.15
% Discount	0.0005139	***	0.0000849	6.05
$ Min. Price	-3.70E-06	**	1.25E-06	-2.95
In-store Purchase	-0.0075032	***	0.0016533	-4.54
Unlimited Use	0.0011744		0.0009001	1.30
"Pay-By-Tweet"	-0.0104336	***	0.0023861	-4.37
Initial Brand Followers	0.0000148	**	5.55E-06	2.67
Social Network Peak Time	0.0041712	**	0.0013765	3.03

Robust standard error are reported
* significant at 10%; ** significant at 5%; *** significant at 1%

ber of followers and show the effect for the same brand and promotional event.

The *statistical significance* of the abnormal returns (AR_{it}) as well as the cumulative abnormal returns (CAR_i) can be tested using the corresponding t-statistics and the corresponding sample standard deviation in a time-series approach taking into account also any cross-sectional dependence. Thanks to the narrow event window, the test-statistic is not very sensitive to the benchmark model of abnormal returns or assumptions about cross-sectional or time-series dependence of abnormal returns [17]. Additionally, the significance of the cumulative abnormal returns for all brands as a group can also be estimated.

Finally, we also conduct a complementary explanatory analysis investigating any potential differences in the estimated (cumulative) abnormal returns of the various firms. In particular, we explain the differences in the magnitudes of the returns across brands and organizations in terms of event characteristics using the following econometric model:

$$CAR_i = y_i = \alpha + \beta \mathbf{x}_i + \gamma w_i + \epsilon_i$$

where $\mathbf{x}_i$ a vector of the relevant characteristics of the social media marketing strategy, including the promotional discount, the minimum amount (USD) a user is required to spend to be eligible for the promotion, whether the promotion is available only for in-store purchases, if the promotion can be claimed multiple times from a single user, and whether the marketing promotion was announced between 12:00pm and 4:00pm, which is the peak of user activity on Twitter platform [34], and w_i the initial size of the brand fan base at the time of announcement of the social media marketing promotion.

5. RESULTS

After learning the model of normal returns for each brand using standard data mining techniques and taking into consideration the observations about the brand itself before the "treatment" as well as the behavior of a set of other time series (that were predictive of the target prior to the intervention) and based on their values on the post-treatment period, we compute the daily abnormal returns and the cumulative abnormal returns for each firm. The daily abnormal returns are computed by subtracting the predicted normal return from the actual return for each day in the event window. Then, the cumulative abnormal returns are measured as the sum of the abnormal returns over the event window. As Table 2 shows, the cumulative abnormal returns are positive and statistically significant (0.0020, $p < 0.001$)[2] and thus provide support for the hypothesis that brands which implement social media marketing strategies incorporating explicit and/or implicit advocacy are more likely to attract more followers and expand their fan base ($H1$), illustrating the long-term effects of such marketing strategies on social media platforms.

The identified effect is also economically significant as it corresponds to an increase of about $5,000$ additional new followers -*beyond the normal expectations*- for an average size brand. Then, as described in the previous section, we also conduct a complementary analysis investigating the differences in the estimated (cumulative) abnormal returns of the various brands and marketing promotions.

Based on the results presented in Table 3, a social media marketing strategy from a brand with a smaller existing fan base is less likely to attract more followers for the participating brand ($H2a$). Besides, a marketing promotion that is announced at the time of peak usage of the social media platform ($H3a$) is more likely to attract more followers for the participating brand. Consistent with prior literature, a promotional event with a larger percentage discount and a smaller required amount as minimum spend is more effective. Additionally, a promotion that is valid online, corresponds to a discount offer and not a specific product purchase (i.e., "Pay-By-Tweet"), or can be claimed multiple times by a customer (i.e., unlimited use) is more likely to attract more followers. Thus, a promotional event with higher expected benefits for customers is more likely to attract more followers for the brand.[2]

[2]The presented results should be interpreted as the percentage increase in the ratio (slope) of the increase in the number of followers, rather than a simple linear or percentage increase in the number of followers.

Table 4. Falsification Check (Pseudo Events)

Variable	Coefficient	Robust Std. Err.	t-statistic
Cumulative Abnormal Returns	-0.0007863	0.00049	-1.59

Robust standard error are reported
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5. Falsification Check (Consumer Search Behavior)

Variable	Coefficient	Robust Std. Err.	t-statistic
Cumulative Abnormal Returns	-0.0159971	0.01247	-1.28

Robust standard error are reported
* significant at 10%; ** significant at 5%; *** significant at 1%

6. ROBUSTNESS CHECKS

Additionally, as robustness checks, we use different dependent variables measuring the performance of social media marketing strategies, such as the percentage change in the number of followers or the number of (new) followers per day, and different numbers (B, F, S) of similar brands as synthetic controls (e.g., $5, \ldots, 50$). Besides, in order to assess the effect of potentially over-fitting our data, apart from using standard approaches such as examining the corresponding learning curves, we also estimate the models using only the performance of brands that belong to the same industry and economic sector or just their average performance, instead of the extended set of control time series that we employed in our main analysis. Similarly, we also evaluate the effect using only similar brands based on the Jaccard coefficient computed using the set of followers of each brand. Furthermore, similar results are observed on average by learning either a pooled regression model of normal performance for all the firms and allowing for firm-specific fixed effects or brand-specific predictive models. Finally, various extra robustness checks are conducted, by varying the length of the event and estimation windows, qualitatively corroborating our empirical findings.

7. FALSIFICATION TESTS

One might think that it is plausible that the previous set of models are simply picking up spurious effects as a result of pure coincidence, a general increase in the corresponding metrics, or other unobserved factors. To assess the possibility that the aforementioned findings are a statistical artifact and the identified positive significant effects were captured by chance or because of other confounding factors, we run a falsification test ("placebo" study) using the same models (in order to maintain consistency) but randomly indicating when each promotional event took place. In particular, we use a standard (pseudo) random number generator in order to create a dummy variable that indicates when each promotional event occurred; when drawing the pseudo event day, we always allow for an estimation window at least 30 days long and a (pseudo) event window that immediately follows this estimation window and does not overlap with the actual event window of the real event. Under this falsification test, since the pseudo event window was not affected by the events of interest, the corresponding variable of interest of abnormal returns should not pick any effect in the falsification models and show that there is no impact. The results of this falsification test are shown in Table 4. We see that, under this check, the corresponding cumulative abnormal returns are negative and *not* statistically significant, indicating that our findings are not a statistical artifact of our specification, but we indeed discovered the actual effects.

In addition, we design and conduct a second set of falsification checks. To assess the possibility that the discovered cumulative abnormal effects are caused by other concurrent events or confounding factors, such as seasonality effects for the specific brand, broader marketing campaigns or offline advertising, we should examine the impact of the specific promotional events under study on other independent marketing objectives that would capture a general increase in consumers' interest for the specific brand. Therefore, we evaluate the potential impact of the promotional events on consumer brand-specific search behavior [16]. In particular, we use as dependent variable the normalized search volume for the brand as captured by Google Trends [12]. Under this falsification test, if the new dependent variable was not affected by the events of interest, the corresponding variable of interest of abnormal returns should not pick any effect in the falsification models and show that there is no impact. The results of this falsification test are shown in Table 5. We see that, under this check, the corresponding cumulative abnormal returns are negative and *not* statistically significant, indicating that our findings are not caused by other confounding factors, but the actual effects we discovered can be attributed to the marketing strategies on the social network.

8. CONCLUSION

Studying the pioneering social media venture of American Express on the social network of Twitter, we contribute to the related literature in social media and e-commerce discovering new rich findings and provide actionable insights with implications for network-based targeting and social media marketing. Based on the proposed combination of data mining and econometric techniques in a causal estimation framework that allows for more accurate counterfactuals, we focus on the long-term business value of this novel application of social media and evaluate the effectiveness of social media marketing strategies towards expanding a brand's social media fan base. Based on our empirical results and a combination of robustness checks and "placebo" studies, we find that social media marketing strategies combined with features enabling implicit or explicit advocacy on social media platforms result in statistically significant positive abnormal returns, in terms of additional new followers for the corresponding brands. In essence, we identify an effective method for brands to attract more followers and expand their social media fan base, which they can later leverage in order to increase awareness, engagement, and word-of-mouth [31]. Besides, we illustrate that social media marketing strategies are more effective for brands that already have a large user base, since they can more effectively propagate their messages. Moreover, we show empirically when consumers are more receptive to such marketing messages on social media and thus suggest specific tactical strategies for firms that would like to capitalize on the real time marketing feature of the Twitter social network. Towards this direction, we precisely quantify the impact of the various marketing promotion characteristics (e.g., discount, price, flexibility, expected benefit) and demonstrate what types of promotions are effective for generating desired long-term effects. We discover that the most important factor in increasing significantly the returns of promotional events is the timing of announcing the promotion and broadcasting the corresponding brand message on social networks. Hence, based on the aforementioned results, one of the tactical strategies with significant monetary implications for brands, organizations, and marketers is that, despite the increased competition for consumers' attention, broadcasting marketing messages on social networks during the time of peak usage maximizes their returns as the messages can be more widely disseminated to the network; even if the effectiveness of a message per user is higher during off-peak hours. Our results highlight the need for future studies to investigate the effect of the timing of marketing messages and public announcements across different social media and social networks. For instance, significantly different tactical strategies might be optimal in the social platform of Facebook where the timelines of the users are algorithmically curated.

A limitation of this study is that selecting a narrow event window we might under-estimate the actual causal effect of the particular social media marketing strategies. Moreover, firms that announced the marketing promotions during other confounding events were excluded from the sample, since the effect could not be untangled from the confounding event and attributed to the particular social media marketing strategy. Such firms might experience even higher abnormal returns because of synergies and complementarity effects. As part of the future work, we plan to study how the effect of social media marketing strategies evolves over time and its decay structure as well as the impact of social media promotional events on consumers' perceptions of participating brands.

9. REFERENCES

[1] American Express Company. $15 Million in Savings Available to Amex Card Members in April 2014. https://twitter.com/AmericanExpress/status/452456770804674560. Posted: 2014-04-05, Accessed: 2015-02-20.

[2] S. Aral. Commentary-identifying social influence: A comment on opinion leadership and social contagion in new product diffusion. *Marketing Science*, 30(2):217–223, 2011.

[3] S. Aral, C. Dellarocas, and D. Godes. Introduction to the special issue-social media and business transformation: A framework for research. *Information Systems Research*, 24(1):3–13, 2013.

[4] J. Y. Bakos. Reducing buyer search costs: Implications for electronic marketplaces. *Management science*, 43(12):1676–1692, 1997.

[5] S. Balasubramanian, R. A. Peterson, and S. L. Jarvenpaa. Exploring the implications of m-commerce for markets and marketing. *Journal of the academy of Marketing Science*, 30(4):348–361, 2002.

[6] E. Brynjolfsson and M. D. Smith. Frictionless commerce? a comparison of internet and conventional retailers. *Management science*, 46(4):563–585, 2000.

[7] J. Chevalier and A. Goolsbee. Measuring prices and price competition online: Amazon. com and barnesandnoble. com. *Quantitative marketing and Economics*, 1(2):203–222, 2003.

[8] J. A. Chevalier and D. Mayzlin. The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, 43(3):345–354, 2006.

[9] C. Dellarocas. The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management science*, 49(10):1407–1424, 2003.

[10] D. DelVecchio, D. H. Henard, and T. H. Freling. The effect of sales promotion on post-promotion brand preference: A meta-analysis. *Journal of Retailing*, 82(3):203–213, 2006.

[11] A. Dickinger and M. Kleijnen. Coupons going wireless: Determinants of consumer intentions to redeem mobile coupons. *Journal of Interactive Marketing*, 22(3):23–39, 2008.

[12] Google Inc. Google Trends. www.google.com/trends/, 2015.

[13] S. Hill and A. D. Benton. Social tv: Linking tv content to buzz and sales. In *Proceedings of the 33th International Conference on Information Systems*, 2012.

[14] S. Hill, F. Provost, and C. Volinsky. Network-based marketing: Identifying likely adopters via consumer networks. *Statistical Science*, pages 256–276, 2006.

[15] A. H. Huang, A. Y. Zang, and R. Zheng. Evidence on the information content of text in analyst reports. *The Accounting Review*, 89(6):2151–2180, 2014.

[16] M. Joo, K. C. Wilbur, B. Cowgill, and Y. Zhu. Television advertising and online search. *Management Science*, 60(1):56–73, 2013.

[17] Y. Konchitchki and D. E. O'Leary. Event study methodologies in information systems research. *International Journal of Accounting Information Systems*, 12(2):99–115, 2011.

[18] R. P. Leone and S. S. Srinivasan. Coupon face value: its impact on coupon redemptions, brand sales, and brand profitability. *Journal of Retailing*, 72(3):273–289, 1996.

[19] LexisNexis Group. LexisNexis U.S. and World News database. http://www.lexisnexis.com/, 2015.

[20] Y. Liu. Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of marketing*, 70(3):74–89, 2006.

[21] X. Luo, J. Zhang, and W. Duan. Social media and firm equity value. *Information Systems Research*, 24(1):146–163, 2013.

[22] A. C. MacKinlay. Event studies in economics and finance. *Journal of economic literature*, pages 13–39, 1997.

[23] D. Molitor, P. Reichhart, and M. Spann. Location-based advertising: what is the value of physical distance on the mobile internet. *Social Science Research Network, Rochester*, 2012.

[24] L. Muchnik, S. Aral, and S. J. Taylor. Social influence bias: A randomized experiment. *Science*, 341(6146):647–651, 2013.

[25] P. Raghubir. Coupons in context: discounting prices or decreasing profits? *Journal of retailing*, 80(1):1–12, 2004.

[26] N. Roztocki and H. R. Weistroffer. Event studies in information systems research: An updated review. In *Proceedings of the 15th Americas Conference on Information Systems*, 2009.

[27] S. Rubel. Trends to watch part ii: Social commerce. *Micro Persuasion vom*, 23:2005, 2006.

[28] N. Shi, Y. Hong, K. Wang, and P. A. Pavlou. Social commerce beyond word of mouth: Role of social distance and social norms in online referral incentive systems. In *Proceedings of the 34th International Conference on Information Systems*, 2012.

[29] I. Simonson, Z. Carmon, and S. O'curry. Experimental evidence on the negative effect of product features and sales promotions on brand choice. *Marketing Science*, 13(1):23–40, 1994.

[30] G. J. Tellis. *Effective advertising: Understanding when, how, and why advertising works*. Sage Publications, 2003.

[31] V. Todri and P. Adamopoulos. Social commerce: An empirical examination of the antecedents and consequences of commerce in social network platforms. In *Proceedings of the 35th International Conference on Information Systems*, 2014.

[32] C. Wang and P. Zhang. The evolution of social commerce: The people, management, technology, and information dimensions. *Communications of the Association for Information Systems*, 31(5):1–23, 2012.

[33] D. Williamson. Advertising on twitter: Unique opportunities outweigh slowing user growth. *March*, 27:2014, 2014.

[34] D. Zarrella. *The social media marketing book*. " O'Reilly Media, Inc.", 2009.

[35] F. Zhu and X. Zhang. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of marketing*, 74(2):133–148, 2010.